UvA-DARE (Digital Academic Repository)

Sequence variation in the hpd gene of noncapsulated *Haemophilus influenzae* isolated from patients with chronic bronchitis

Duim, B.; Ruiter, P.; Bowler, L.; Dankert, J.; van Alphen, A.J.W.

DOI

[10.1016/S0378-1119\(97\)00030-9](https://doi.org/10.1016/S0378-1119(97)00030-9)

Publication date

1997

Published in

Gene

[Link to publication](#)

Citation for published version (APA):

Duim, B., Ruiter, P., Bowler, L., Dankert, J., & van Alphen, A. J. W. (1997). Sequence variation in the hpd gene of noncapsulated *Haemophilus influenzae* isolated from patients with chronic bronchitis. *Gene*, 191, 57-60. [https://doi.org/10.1016/S0378-1119\(97\)00030-9](https://doi.org/10.1016/S0378-1119(97)00030-9)

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Sequence variation in the *hpd* gene of nonencapsulated *Haemophilus influenzae* isolated from patients with chronic bronchitis

Birgitta Duim¹, Paula Ruiter, Lucas D. Bowler², Jacob Dankert, Loek van Alphen *

Department of Medical Microbiology, Room L1-162, University of Amsterdam, Academic Medical Center, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands

Received 19 August 1996; received in revised form 12 December 1996; accepted 13 December 1996; Received by A. Bernardi

Abstract

The molecular diversity of protein D of nonencapsulated *Haemophilus influenzae* strains isolated from persistently infected patients with chronic bronchitis was studied by sequencing the *hpd* gene of four independently obtained isolates. The nucleotide (nt) sequences of the *hpd* genes of two strains were identical. The other two *hpd* sequences showed nt substitutions which were mostly synonymous. As a consequence the deduced amino acid (aa) sequences differed from the consensus sequence only by a few aa. No changes in the *hpd* genes were observed among the four variants of the four strains persisting in chronic bronchitis patients for 9, 11, 8 and 3 months, respectively, although variation in their major outer membrane proteins P2 and P5 occurred. We conclude that the *hpd* gene is conserved during chronic infections of nonencapsulated *H. influenzae*. © 1997 Elsevier Science B.V.

Keywords: PCR; DNA sequencing; Protein D; Microbial pathogenesis

1. Introduction

Nonencapsulated *Haemophilus influenzae* causes acute respiratory tract infections and persistent infections in the lower respiratory tract of patients with chronic bronchitis and cystic fibrosis (van Alphen, 1992; Murphy and Sethi, 1992; van Alphen et al., 1995). The major outer membrane proteins (MOMPs) P2 and P5 of nonencapsulated *H. influenzae* are important antigens. Since the composition of MOMPs P2 and P5 is very heterogeneous and shows antigenic drift during persistent infections in patients with chronic bronchitis (Duim et al., 1994), the application of these MOMPs as vaccine components seems to be restricted.

Protein D is a 42 kDa lipoprotein of *H. influenzae*

(Janson et al., 1992). Protein D of nonencapsulated *H. influenzae* and *H. influenzae* type b was shown to be antigenically conserved (Akkoyunlu et al., 1991), and sequence analysis of the *hpd* gene encoding protein D showed a limited diversity (Janson et al., 1993; Song et al., 1995). A protein D-deficient mutant of *H. influenzae* was strongly impaired in inducing experimental otitis media in rats compared to the wild-type strain (Janson et al., 1994), indicating that protein D is involved in the pathogenesis of this infection. Although it has been suggested that protein D possesses an IgD binding site (Janson et al., 1991; Akkoyunlu et al., 1991), it shows approx. 67% sequence homology with the glycerophosphodiester phosphodiesterase (GlpQ) of *Escherichia coli*, which implies that protein D may have a major periplasmic domain (Munson and Sasaki, 1993).

The antigenic conservation of protein D and the role of this protein in the onset of *H. influenzae* infection suggest that protein D is a candidate antigen for a vaccine to prevent nonencapsulated *H. influenzae* infections. To analyze whether protein D is also conserved during persistent infections we analyzed the *hpd* gene of *H. influenzae* strains isolated from patients with chronic bronchitis and compared the conservation of the *hpd* gene of *H. influenzae* strains varying in their MOMPs P2 and P5 during such persistent infections.

* Corresponding author. Tel. +31 20 5664862; Fax +31 20 6979271; e-mail: a.j.vanalphen@amc.uva.nl

¹ Present address: Institute of Animal Science and Health, Department of Bacteriology, Lelystad, The Netherlands.

² Present address: Department of Molecular Microbiology, University of Sussex, Brighton BN1 9QG, UK.

Abbreviations: aa, amino acid(s); bp, base pair(s); *hpd*, gene encoding protein D; IgD, immunoglobulin D; kDa, kilodalton(s); MOMP, major outer membrane protein; nt, nucleotide(s); PAGE, polyacrylamide-gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate.

2. Experimental and discussion

2.1. Amplification and sequence analysis of the *hpd* genes

Nonencapsulated *H. influenzae* strains A850053, A850047, A850079, and A850048 were independent isolates from sputum samples of four patients with chronic bronchitis (Groeneveld et al., 1988). The *hpd* genes of these *H. influenzae* strains were obtained with PCR. Using the primers listed in Table 1, a product of approx. 1029 bp was amplified comprising 94% of the ORF of the *hpd* gene sequence (Janson et al., 1993). Of the *hpd* genes the 57 nt at the 5' end encoding the signal sequence of protein D were not determined. This region has been shown to be conserved (Song et al., 1995). Also as this region is not part of the processed form of protein D, nt substitutions may be irrelevant for the functioning of protein D.

Alignment of the *hpd* sequences of the four independent *H. influenzae* strains revealed only minor nt differences (Fig. 1). Their *hpd* sequences were 96.3% identical, two strains possessed identical *hpd* sequences, the other two strains differed by 14 and 32 nt substitutions, resulting in 2 and 5 aa changes, respectively. As the C-terminal part of protein D is suggested to contain the surface-exposed IgD-binding region, this region may experience a selective pressure caused by the immune system of the host and could be more variable than the N-terminal part. However, the aa differences were randomly distributed along the gene. The *hpd* gene sequences were very similar to those reported by Song et al. (1995), indicating that the nt sequences of the *hpd* genes and consequently the deduced aa sequences, of different nonencapsulated *H. influenzae*, including *H. influenzae* type b are highly homologous independent of their source.

Next, the *hpd* gene of four *H. influenzae* variants A860509, A860501, A860503 and A850080 of the four

strains were analyzed in the same way. From sputum samples of the same four patients with chronic bronchitis from whom the strains A850053, A850047, A850079 and A850048, respectively, were recovered (Groeneveld et al., 1988; Duim et al., 1994) variants were isolated that were genotypically identical but showed differences in mobility of their MOMP P2 and/or P5 proteins. The genotyping was performed by analysis of DNA fingerprints after digestion with two restriction enzymes (Groeneveld et al., 1988) and by random amplified polymorphic DNA (RAPD) typing (van Belkum et al., 1994). Time intervals between isolation of strains and their variants were 9, 11, 8 and 3 months, respectively. Although genetic variation was observed in their MOMP P2 and P5 genes the *H. influenzae* MOMP variants showed no differences in their *hpd* coding sequences (Duum et al., 1994, 1996).

2.2. The nature of the *hpd* sequence variation

For occurrence of mutations an equal probability of synonymous (silent) or nonsynonymous substitutions (aa altering) is expected. The number of changes is dependent on the number of synonymous and nonsynonymous sites present in a given nt sequence (Sharp, 1991). The numbers of synonymous and nonsynonymous nt substitutions was calculated, using the method of Nei and Gojobori (1986). The differences in the *hpd* genes of four distinct *H. influenzae* strains were mostly synonymous substitutions instead of nonsynonymous substitutions (Fig. 1). Such a result is normally found in bacterial genes that suffer from some functional or structural constraints (Sharp, 1991).

2.3. Secondary protein structure prediction

Using the method of Garnier et al. (1987) for secondary protein structure prediction based on the deduced

Table 1
Oligonucleotide primers used for amplification and sequencing of the *hpd* gene

Name	Position in the <i>hpd</i> gene ^a	Sequence (5'-3') ^b
hpd1	58–78	<i>TGT AAA ACG ACG GCC AGT</i> AGC AGC CAT TCA TCA AATATG
hpd2	385–403	<i>TGT AAA ACG ACG GCC AGT</i> GAT GGC AAA CAA GCG CAA G
hpd1r	1086–1068	<i>CAG GAA ACA GCT ATG ACC</i> TCC TTT TAA GAA TCC CAC G
hpd2r	741–721	<i>CAG GAA ACA GCT ATG ACC</i> CTT TCC AAT CTG TAT AAG C

^aNumbering derived from the published sequence of the *hpd* gene of *H. influenzae* type b (Janson et al., 1993).

^bThe –21m13 sequence (TGT AAA ACG ACG GCC AGT) and the M13 reverse sequence (CAG GAA ACA GCT ATG ACC), added at the 5' end for direct sequencing of the PCR products with fluorescently labeled M13 primers (Duum et al., 1993) are indicated in italics. Overlapping templates of the P5 gene were amplified by PCR. A cell lysate of *H. influenzae* was used as source of chromosomal DNA (Duum et al., 1993). The PCR was carried out in a total volume of 100 µl, containing 10 µl chromosomal DNA preparation, 100 pmol of each primer, 200 µm of each deoxynucleotide (dATP, dCTP, dGTP, dTTP), 0.01% gelatin, 2.5 mM MgCl₂, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, and 2.5 U of Amplitaq DNA polymerase (Perkin Elmer Cetus). Following initial denaturation at 95°C for 5 min, 35 amplification cycles were performed in a Trio Thermoblock (Biometra). Each cycle comprised 1 min at 95°C, 1 min at 50°C, and 1 min 72°C. The PCR was finished with a final incubation at 72°C for 5 min. PCR products were analyzed on a 1% agarose gel using a Tris-acetate-EDTA buffer system (Maniatis et al., 1982). Relevant products were excised from the gel and purified by QIAEX according to the instructions supplied by QIAGEN, Chatsworth, CA. The DNA concentration was adjusted to approx. 1 ng/µl in distilled water for sequencing.

A850053	<u>AGCAGCCATTTCATCAAAATATGGCGAATACCCAAATGAAATCAGACAAAAT</u>	50
A850047		50
A850079	A.....T.....	50
A850048		50
	S S H S S N M A N T Q M K S D K	16
		
A850053	CATTATTGCTCACCGTGGTGCTAGCGGTTATTTACCAGAGCATACGTTAG	100
A850047		100
A850079	...C.....T.....	100
A850048		100
	I I I A H R G A S G Y L P E H T L	33
		
A850053	AATCTAAAGCACTTGCGTTTGCAACACGCAGATTATTTAGAGCAAGAT	150
A850047	G.....T.....	150
A850079	G.....T.....C.....	150
A850048		150
	E S K A L A F A Q H A D Y L E Q D	50
		
A850053	TTAGCAATGACTAAGGATGGTCGTTTAGTGGTTATTCACGATCACTTTTT	200
A850047		200
A850079		200
A850048		200
	L A M T K D G R L V V I H D H F L	66
		
A850053	AGATGGCTTGACTGATGTTGCGAAAAAATCCACATCGTCACCGTAAAG	250
A850047	T.....	250
A850079		250
A850048		250
	D G L T D V A K K F P H R H R K	83
		
A850053	ATGGTCGTTACTATGTCATCGACTTTACCTTAAAAGAAATTCAAAGTTTA	300
A850047		300
A850079	...C.....	300
A850048		300
	D G R Y Y V I D F T L K E I Q S L	100
		
A850053	GAAATGACTGAAAACCTTTGAAACCAAGACGGCAAACAAGCGCAAGTTTA	350
A850047	A.....T.....	350
A850079	A.....T.....GA.A.....	350
A850048		350
	E M T E N F E T K D G K Q A Q V	116
		
A850053	TCCTAATCGTTTCCCACTTTGGAAATCACATTTAGAAATTCACACCTTGG	400
A850047	T.....	400
A850079	C.....C.....CC.T.....	400
A850048		400
	Y P N R F P L W K S H F R I H T F	133
		
A850053	AAGATGAAATTGAGTTTATCCAAGGCTTAGAAAAATCGACTGGCAGAAAA	450
A850047	A.....T.....C.....A.....	450
A850079	A.....T.....T.....G.....A.....	450
A850048		450
	E D E I E F I Q G L E K S T G R K	150
		
A850053	GTAGGGATTTATCCAGAAATCAAAGCACCTTGGTTCCACCATCAAAATGG	500
A850047		500
A850079		500
A850048		500
	V G I Y P E I K A P W F H H Q N G	166
		
A850053	CAAAGATATTGCAGCTGAAACGCTCAAAGTGTTAAAAAAATATGGCTATG	550
A850047	T.....T.....	550
A850079		550
A850048		550
	K D I A A E T L K V L K K Y G Y	183
		

Fig. 1. Comparison between *hpd* gene sequence and deduced aa sequence of four distinct nonencapsulated *H. influenzae* strains. PCR fragments obtained in two independent PCR reactions were sequenced in both directions using the *Taq* Dye Primer sequencing system with fluorescent dye-labeled –21m13 and M13 reverse sequence primers (Applied Biosystems). The primer sequences are underlined. DNA analysis was performed with an automated fluorescent DNA sequencer, Model 370A, Applied Biosystems. The derived sequences were analyzed using computer programs included in the program package Pcgene (Intelligent, Inc., 1991). The nt sequences were aligned with the program CLUSTAL using the method developed by Higgins and Sharp (1988). The nt sequences will appear in the EMBL/GenBank/DBJ Nucleotide Sequence Databases under the accession Nos. X90493, X90489, X90491, X90495.

aa sequences, no regions of high surface probability were apparent. This is consistent with the lipoprotein character of protein D. Protein D was also indicated as a IgD-binding protein. As described by Janson et al. (1991) the basic structure of protein D differs from all other characterized Ig-binding proteins.

3. Conclusions

- (1) The *hpd* gene encoding protein D of nonencapsulated *H. influenzae* is conserved. Analysis of the minor changes which were found, and the lack of potentially surface-exposed regions of protein D, indicate that variation of the *hpd* gene sequence by selective pressure most likely does not occur, but rather shows some structural constraints.
- (2) Although the sequence of protein D of nonencapsulated *H. influenzae* is conserved during persistent infections of *H. influenzae* in patients with chronic bronchitis, the usefulness of protein D in a potential *H. influenzae* vaccine can be discussed. As protein D is probably not localized at the bacterial surface the development of a protective immune response against protein D may be hampered. Further research is needed to confirm the exact function and antigenicity of protein D and the potential as a vaccine antigen.

References

- Akkoyunlu, M., Ruan, M. and Forsgren, A. (1991) Distribution of protein D, an immunoglobulin D-binding protein, in *Haemophilus* strains. *Infect. Immun.* 59, 1231–1238.
- Duim, B., Dankert, J., Jansen, H.M. and van Alphen, L. (1993) Genetic analysis of the diversity in outer membrane protein P2 of non-encapsulated *Haemophilus influenzae*. *Microb. Pathog.* 14, 451–462.
- Duim, B., van Alphen, L., Eijk, P., Jansen, H.M. and Dankert, J. (1994) Antigenic drift of non-encapsulated *Haemophilus influenzae* major outer membrane protein P2 in patients with chronic bronchitis is caused by point mutations. *Mol. Microbiol.* 11, 1181–1189.
- Duim, B., Bowler, L.D., Jansen, H.M., Dankert, J. and van Alphen, L. (1996) Molecular variation in the MOMP P5 gene of non-encapsulated *Haemophilus influenzae* during chronic infection. *Infect. Immun.* submitted.
- Garnier, J., Osguthorpe, D.J. and Robson, D. (1987) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120, 97–120.
- Groeneveld, K., van Alphen, L., Eijk, P.P., Jansen, H.M. and Zanen, H.C. (1988) Changes in outer membrane proteins of nontypeable *Haemophilus influenzae* in patients with chronic obstructive pulmonary disease. *J. Infect. Dis.* 158, 360–365.
- Higgins, D.G. and Sharp, P.M. (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* 73, 237–244.
- Janson, H., Heden, L.-O., Grubb, A., Ruan, M. and Forsgren, A. (1991) Protein D, an immunoglobulin D-binding protein of *Haemophilus influenzae*: cloning, nucleotide sequence, and expression in *Escherichia coli*. *Infect. Immun.* 59, 119–125.
- Janson, H., Heden, L.-O. and Forsgren, A. (1992) Protein D, the immunoglobulin D-binding protein of *Haemophilus influenzae*, is a lipoprotein. *Infect. Immun.* 60, 1336–1342.
- Janson, H., Ruan, M. and Forsgren, A. (1993) Limited diversity of the protein D gene (*hpd*) among encapsulated and nonencapsulated *Haemophilus influenzae* strains. *Infect. Immun.* 61, 4546–4552.
- Janson, H., Melhus, A., Hermansson, A. and Forsgren, A. (1994) Protein D, the glycerophosphodiester phosphodiesterase from *Haemophilus influenzae* with affinity for human immunoglobulin D, influences virulence in a rat otitis media model. *Infect. Immun.* 62, 4848–4854.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. CSLH Press, Cold Spring Harbor, NY.
- Munson, R.S. Jr. and Sasaki, K. (1993) Protein D, a putative immunoglobulin D-binding protein produced by *Haemophilus influenzae*, is glycerophosphodiester phosphodiesterase. *J. Bacteriol.* 175, 4569–4571.
- Murphy, T.F. and Sethi, S. (1992) Bacterial infection in chronic obstructive pulmonary disease. *Am. Rev. Respir. Dis.* 146, 1067–1083.
- Nei, M. and Gojobori, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3, 290–300.
- Sharp, P.M. (1991) Determinants of DNA sequence divergence between *Escherichia coli* and *Salmonella typhimurium*: codon usage, map position, and concerted evolution. *J. Mol. Biol.* 33, 23–33.
- Song, X.-M., Forsgren, A. and Janson, H. (1995) The gene encoding protein D (*hpd*) is highly conserved among *Haemophilus influenzae* type b and nontypeable strains. *Infect. Immun.* 63, 696–699.
- van Alphen, L. (1992) Epidemiology and prevention of respiratory tract infections due to nonencapsulated *Haemophilus influenzae*. *J. Infect. Dis.* 165, S177–S180.
- van Alphen, L., Jansen, H.M. and Dankert, J. (1995) Virulence factors in the colonization and persistence of bacteria of the lower airways. *Am. J. Respir. Crit. Care Med.* 151, 2094–2100.
- van Belkum, A., Duim, B., Regelink, A., Möller, L., Quint, W. and van Alphen, L. (1994) Genomic DNA fingerprinting of clinical *Haemophilus influenzae* isolates by polymerase chain reaction amplification: comparison with major outer-membrane protein and restriction fragment length polymorphism analysis. *J. Med. Microbiol.* 41, 63–68.